FINAL REPORT

# Ridership Technical Advisory Panel Review of the California High-Speed Rail Ridership and Revenue Forecasting Process

Findings and Recommendations from the Nov. 2013-Jan. 2014 Review Period

March 28, 2014

The Ridership Technical Advisory Panel (RTAP) held a formal meeting on January 16-17, 2014 at the Parsons Brinckerhoff offices in San Francisco. It was the first meeting as the rebranded RTAP, after meeting ten times as an independent peer review panel. This report covers the Panel's review of progress from November 2013 through the end of January, 2014. The panelists remain from the independent peer review panel:

- Frank S. Koppelman, PhD, Professor Emeritus of Civil Engineering, Northwestern University (chair)
- Kay W. Axhausen, Dr.Ing., Professor, Institute for Transport Planning and Systems, ETH Zurich (Swiss Federal Institute of Technology Zurich)
- Eric Miller, PhD, Professor, Department of Civil Engineering, University of Toronto
- David Ory, PhD, Principal Planner/Analyst, Metropolitan Transportation Commission
- Kenneth A. Small, PhD, Professor Emeritus, Department of Economics, University of California-Irvine

All panelists were present in person for the meeting. Rick Donnelly, PhD, of Parsons Brinckerhoff (PB) served as facilitator and recorder for the Panel. Thierry Prate of PB was invited to attend the meeting as a representative of the program management team. The meeting was otherwise closed to non-members. The Panel briefed staff of Cambridge Systematics (CS), the demand consultant, by teleconference of the findings summarized in this report at the end of the meeting. Their initial feedback is incorporated below as appropriate, while in other cases their responses clarified issues raised by the Panel or obviated the need for further discussion.

# 1 Re-designation as the Ridership Technical Advisory Panel

The Panel operated as an independent peer review panel from its appointment in late 2010 until the end of 2013. They reported directly to the executive director of the California High-Speed Rail Authority and its Board of Directors. The new designation stems from the Authority's reliance on the Panel for advice rather than just reviews of completed work. The mission, independence, composition, and reporting of the Panel remains unchanged in all other respects.

# 2 Version 2 Model Development

CS provided the Panel with a 91-page briefing book prior to the meeting. It described their progress to date on the Version 2 model calibration and validation work. This work had been completed in three months since the Panel's previous meeting. In general the Panel was pleased with the content and results presented, as well as the impressive amount of progress that had been made since the last meeting. This section, except for Subsection 2.2, is based upon that briefing book.

## 2.1 Model Calibration

The calibration process followed standard practices, making the best use of existing data, showing appropriate restraint in adding constants, and avoiding over-fitting the data. The Panel was pleased that the interchange-specific constants used in the Version 1 destination choice models have been removed. It was likewise encouraging that the trip frequency rates did not require large adjustments during the calibration process.

The size of the constants in the main mode choice model was discussed at length. The fact that all the modal constants are large and negative suggests that the model is unable to fully explain why automobile travel is so dominant. This makes the development, testing, and interpretation of the high-speed rail (HSR) constant especially difficult, as discussed further below. This suggests the need for a Version 3 model designed to incorporate important attributes of long-distance travel (e.g., party size, duration) and household characteristics (e.g., size by age group, auto ownership and residential density).

The large modal constants and the large changes in them during calibration also make a strong case for ongoing data collection. Long-distance travel is an infrequent, and thus relatively expensive to survey, behavior. The Panel believes that if the Authority spends in excess of $1 million per year on ridership forecasting, a substantial amount of that (a minimum of 10 percent, but more likely in the range of 20-25 percent) should be devoted to data collection. Better information on access, egress, and main mode choice for long-distance trips will be required, as will be data on time of day choice, access and egress station choice, and trip duration. These more detailed data will be needed to support the more detailed models the Authority will need in the future. David Ory will reach out to the Authority on behalf of the Panel to discuss coordination of data collection efforts with the MPOs.

The Panel has asked to see the pre- and post-calibration goodness of fit measures versus observed summaries in order to better assess the impact of the calibration, as well as the possibility that smaller parameter values might be adequate. It raised two specific questions about the calibration process:

- Might smaller alternative-specific constants work almost as well in the main mode choice model? For example, instead of a value of (say) -12.0, might a value of -6.0 work almost as well? More generally, might it be better to place a limit on the size of the constants, accepting less precision in matching calibration targets in order to increase the model's ability to make use of observed data?
- What would happen in the destination choice model if the destination-specific constants, which were not in the estimated model but introduced during calibration, were omitted, or constrained to be of smaller magnitude? We are simply seeking this as information to help judge the role such constants are playing in reaching calibration targets.

## 2.2 HSR Constant Specification

This subsection is based upon work prepared by CS that describes how the constants in the mode choices were calibrated, with a focus on the HSR constant. The Panel paid special attention to this process, as the calibrated HSR constant is very important to the model outcomes.

Four air constants had previously been calculated based upon the size (major versus minor) of the originating and terminating airport. The Panel had suggested either (1) weighting the four air constants before combining with other information, or (2) estimating four separate HSR constants (based on Stated Preference survey responses) for a similar breakdown of major versus minor stations, rather than just one as now. CS elected to use method (1), obviating the need for re-estimating the model. The Panel finds this decision acceptable, although it notes that it may obscure differences between the different markets within California, which may in turn affect how well the model can track the different results from different phases of the project. Thus, an

alternative approach worth exploring would be to estimate HSR SP constants that are differentiated across geographical markets in a manner similar to that observed for air travel, and to use these rather than a single HSR SP constant as the starting point of the adjustments in calibration that have been described by CS.

A key fact in the calibration is that the non-auto mode constants, estimated from survey data, were adjusted downward during calibration, or order to match observed aggregate mode use. The question is, how one should adjust the HSR constant correspondingly? This is discussed further in Section 3.2.

Additional information or interpretations that would be helpful include:

- A direct comparison of the final HSR constants to those initially estimated.
- An explanation of how CS would relate the HSR constants to those used in the risk analyses (see Section 3.2).
- Consider whether any of the unexplained variation captured in the constants can be treated in an analogous manner to wait and terminal times for other modes of transport (e.g., air, conventional rail). What characteristics contribute to the remaining unexplained variation (e.g., seating comfort, ability to move around)? We urge CS to discuss such possibilities in the report, in order to help readers interpret what behavior is being represented by the mode-specific constants found during estimation and calibration. In the longer term, we urge that additional survey work, in particular Stated Preference (SP) surveys, be considered in order to ascertain the amount that specific attributes may contribute to people's preferences for automobile for many trips. Some questions already incorporated into the 2013-2014 RP-SP survey now underway may provide a start on this agenda.

## 2.3 Year 2010 Model Validation

The results of the 2010 model validation look reasonable, although additional detail is desired in some cases, especially for the all-important intercity air versus HSR flows in the Bay-to-Basin market.

It appears that the model is substantially over- and under-predicting air trips in specific markets. The report notes some possible reasons why the calibration targets may not always be accurate, for example by including trips made by travelers not residing in California. The report needs a more quantitative assessment of the plausible magnitudes of differences between the calibration targets and the forecasts generated by the model.

The model under-predicts intercity rail flows by line. The Panel concludes that this is likely due to the inclusion of trips less than 50 miles in length in the counts. If this assumption is correct, then the shapes of the predicted curves generally look reasonable, and their magnitudes match observed flows well for longer trips.

## 2.4 Year 2000 Model Validation

The model consistently under-predicts year 2000 air travel, an outcome the Panel explored in depth. The CS report suggests that the September 2001 terrorist attacks have fundamentally changed the patterns we see today, reducing air travel relative to 2000 even after adjusting for

observable factors accounted for by the model. The Panel questions this explanation. The enduring legacy of the terrorist attacks is arguably not fear of flying, but rather increases in terminal time and inconvenience; therefore an analysis of the effects of increased terminal times, based on other research and the literature, might clarify whether the report's explanation is plausible.

The Panel instead suspects that the main reason for under-prediction of year 2000 travel is that the model is not adequately capturing the effects of macroeconomic trends. In particular, it does not respond much to changes in employment as a fraction of population. This suggests that the current model, being fit on 2010 data, a year affected by a severe recent recession, may be overly pessimistic about the future, just as the earlier model fit on year 2000 data was overly optimistic before it was adjusted for 2010 conditions in its application to the 2012 Business Plan.

It proved difficult to assess cases where different markets (trips less than versus more than 50 miles in some cases) are mixed together. These should be reported separately as long as the 50-mile boundary is used in the model.

## 2.5 Sensitivity Analyses

CS tested the Version 2 system by running a scenario based on the Northeast Corridor (NEC). A similar "NEC-like" comparison had been carried out with the Version 1.1 system, and the analyses of that test lent credibility to the model. At the level of total ridership the results presented for this test of Version 2 appear to be valid, and the reported elasticities appear to be within a reasonable range. This is an encouraging outcome, especially given that California has nothing like the NEC's extensive regional service. Incomplete information makes it difficult to further assess the test results. For example, in their test of a case when the California HSR service levels approximating those of the NEC, a comparison of only the business/commute trip purpose is shown, and only for three region pairs. The Panel would like to see the entire picture in order to make a full and informed assessment. The reporting of additional elasticities besides mode choice (e.g., trip frequency versus changes in population or employment) would be helpful. An analysis of induced travel would also be useful.

## 2.6 Comparison to the Version 1 Forecasts

CS compared the results of the preliminary Version 2 forecasts to those obtained using Version 1 of the model, as the Panel had requested earlier. The latter did not include the Version 1.1 adjustments carried out for the 2012 BP update, and thus represents the scenario used as the "high" forecasts for the 2012 BP. The Panel is not sure that this represents the most useful comparison, as the Authority and the public will presumably want to understand the reason for differences between the new forecasts and those upon which decisions were based made in 2012 (i.e., the "medium" Version 1 forecast).

Moreover, the comparison is somewhat complicated by the use of different inputs in some cases. The Version 2 skim matrices were used with both model systems in order to reduce the amount of effort required to re-create the original Version 1 data. This makes it difficult to determine whether the changes observed are due to differences in the model version or to differences in the skims.

As with the sensitivity testing, additional detailed information is needed in order to fully appraise how the two models compare. The Panel would like to see the actual distributions rather than data grouped by broad distance ranges: grouping the trips into three ranges (<50 miles, 50-99.9 miles, and ≥100 miles) is too coarse to fully understand model performance. HSR mode shares by distance and loading charts (boardings and alightings by station) would be useful comparisons as well.

One key reported finding was that the Version 2 system forecasted a different mix of trip distance in 2030 than obtained with the Version 1 system. Version 2 predicts fewer HSR trips for business/commute purposes but more for other purposes. Substantially more overall trips were forecasted in the 50-99.9 mile range, and fewer trips beyond 100 miles. It is important to report share of the total trips occur entirely within the Metropolitan Transportation Commission (MTC) or Southern Association of Governments (SCAG) areas, for some of the difference might be attributable to special characteristics of trips within those regions in the Version 1 and 2 systems. It would similarly be useful to know how many of the 50-99.9 mile trips do not take place within those regions, for such trips would not have been included in the Version 1 system. A station pair matrix comparison would help inform such reviews.

# 3 Risk Analysis Methodology

The Panel reviewed in depth the draft memorandum, dated January 18, 2014, entitled "California High Speed Rail Ridership and Revenue Forecasts Methods and Forecasts including Risk Analysis." The Panel regards this as a key document for the forecasting efforts, as it provides not only results, but also a framework for interpreting results. The Panel therefore paid considerable attention both to the substantive methodology and the way it will need to be presented as part of the overall documentation of the forecasting effort.

As a technical memorandum, it is competent and sophisticated. The Panel is not aware of a comparable risk analysis performed for a public transportation infrastructure project in the USA. There are several reasons for this assessment:

- The risk analysis recognizes formally that forecasts are inherently stochastic in nature, as opposed to the more common practice of generating a single forecast or a small number of alternate forecasts.
- The analysis gives users a tool to assess quantitatively how uncertainty in inputs and model parameters translate into uncertainty in the model forecasts.
- The analysis facilitates testing of specific sources of uncertainty in terms of what effects each has on revenue predictions.
- The analysis is compatible with the type of risk analyses commonly used for private investment, as it illuminates the effects of multiple sources of uncertainty on any given desired benchmark results, such as total riders or total revenue.
- The specific assumptions used are spelled out in a transparent way, making it easy to assess their realism and, if desired, to substitute alternative assumptions.

As a report on the risk analysis findings, however, the document is lacking in interpretation thus far. Only a single paragraph is devoted to interpreting and discussing the key finding of the

importance of the HSR constant. The consultant should provide readers with a more detailed assessment of how to interpret the results. A technical critique, thoughts on interpretation, and some conclusions about the overall forecasting effort for California HSR are presented in the following sections.

## 3.1 Technical Critique

The risk analysis consists of three steps:

1. Running the forecasting model under many alternate combinations of input assumptions.
2. Fitting a "revenue regression model" (henceforth "regression") to approximate the forecasting model's response to the inputs.
3. Performing a Monte Carlo analysis of risk by using the regression to predict outcomes under a very large number of input assumptions chosen as random draws from specified distributions.

The Panel has the following concerns about the specific choices made in these steps:

- The range and distribution of values deemed reasonable for inputs to take, which is relevant to the Monte Carlo analyses, are discussed as part of the first step of the analysis whose sole purpose is to generate data for fitting the regression. However, the requirements for fitting the regression are not necessarily the same as those for the Monte Carlo experiments. In particular, it might be desirable to choose wider ranges of inputs for fitting purposes than would be used for Monte Carlo, in order to more easily discern the exact form the regression takes.
- As a result of the previous point, the section titled, "Range of Risk Factor Values and Distributions," would be better divided into two parts. The discussion of ranges for forecasting should be discussed after the regression has been described.
- The technical report needs to explain how the regression model was econometrically estimated, which we understand from replies to our questions to be ordinary least squares (OLS) after taking logarithms of both sides of the equation.
- The Panel is pleased that care was taken to include runs in which more than one input was varied simultaneously, in order to capture possible interactions among the inputs. However, the regression model shown does not include any such interactions, but rather explains the logarithm of revenue as a linear function of input variables. We understand from later discussions that such interactions were tested and rejected, and urge that this be stated in the technical report.
- Similarly, it is commendable that intermediate values of inputs were used in order to capture non-linear responses. However, the form of the regression function rules out such non-linear responses. The Panel cannot tell to what extent they were tested, although we understand from discussions that some other functional forms were tested.
- There is no reason why all of the right-hand-side variables must be entered in the same form (as levels in the chosen regression). The form for each of them can be chosen based on the consultants' knowledge of the properties of the forecasting model. We are aware

that some other forms were experimented with, and urge that this information be included in the technical report.

- The name “overall growth” is confusing and misleading as to the nature of the first variable. As we understand it, the variable is simply population, scaled by 2010 population. It is not a growth rate of any kind. This is good, because population, not its growth rate, can logically be expected to influence revenues.
- It would be helpful to show the standard error of regression in addition to the R-square statistic. This would indicate how much deviation in revenues arises from imperfect fit of the regression model to the actual forecasting model runs. Note this is different than the standard errors of the individual coefficients.
- It appears that the regressions for the years 2029 and 2040 use exactly the same rail system (Phase 1 Blended). Thus, perhaps the same inputs were used except for those included as variables in the model. If so, the model runs for those two years could be pooled and a single regression estimated to apply to both of them. We note that in fact the two regressions reported for these two years are nearly identical. Doing this may facilitate the ability to measure nonlinear responses or interactions between input variables if they exist.

## 3.2 Interpretation of Results

The results of the risk analysis suggest that the HSR constant is the dominant source of uncertainty in the forecasts. This is an important finding, meriting much more discussion in the documentation. It would be valuable to describe why this occurs, how certain it is, and what it means for understanding forecasting. Such a discussion should address several points:

- The report appears to suggest that 90 percent of the uncertainty in the HSR forecasts is due to “the attractiveness of HSR within the California transportation environment” (page 25). Is CS making this claim? An alternate interpretation of its findings are as follows:
    - The model specification has a limited understanding of the motivations of mode choice behavior.
    - Large mode-specific constants are required to overcome this lack of understanding when the model is calibrated to actual aggregate travel statistics.
    - The range selected for the HSR mode-specific constant is logically large (because the air and conventional rail constants are large in magnitude).
    - Thus, the range of HSR constants selected for the risk analysis dominate the results.

  If one were to follow this line of thinking, the size of the air and conventional rail constants (which then inform the HSR constant) significantly muddy the risk assessment waters, making it difficult to say whether the uncertainty is due to limited knowledge of HSR attractiveness or to limited knowledge of the determinants of existing air and rail usage.
- Building from the above point, the central problem represented by large model constants is not only a lack of knowledge of travel behavior regarding HSR, but also lack of sufficient knowledge of the determinants of use of travel by conventional rail (CVR) and

by air. The HSR constant is a function of the calibrated CVR and air constants. Two notable facts from the section, "Determining HSR constants," are that many mode-specific constants in the model are large (i.e., have absolute values substantially greater than one) and that they required substantial adjustment in calibration. Even before considering HSR, this suggests that there are important elements of travel behavior left unexplained by the exogenous model variables and captured instead by these constants. . To really understand the meaning of the risk analysis one needs a discussion about what some of those behavioral elements might be, as noted in Section 2.2 above.

- The discussion in this document needs to be connected to that in "Determining HSR constants" (the document discussed in Section 2.2 above), describing alternate ways to set the HSR constant based on the estimation and calibration of the overall model. The memo on risk analysis simply suggests a range of uncertainty for this constant, and reports on how alternate values within that range affect the forecasts. Because these effects are so large more attention is warranted concerning how the range is chosen. Is it supported by evidence, or by expert judgment? The lack of discussion implies the latter, but we believe the former is better. In particular, the section, "Determining HSR Constants," provides three apparently objective methods that yield quite different calibrated HSR constants. Yet only one is used here. We would expect the variation from these three alternate methods to play a key role in choosing the hypothesized distribution of the HSR constant for Monte Carlo analysis.
- The previous points suggest the value of a fourth method for choosing the HSR constant. Expert judgment about the likely sources of behavior represented by the modal constants for CVR and air, and to what extent those sources apply to HSR, will be useful. While we recognize the potential for criticism of any application of expert judgment, we believe such judgment is inherent in any forecasting exercise, including the three alternate methods of choosing an HSR constant already described. Thus, it would be appropriate to include this fourth method, along with the other three, at least to discuss the impact of all four methods on the resulting uncertainty in the HSR constant.
- Some interpretations of the role of modal constants may suggest policies for the Authority to consider. For example, if much of the unknown behavioral sources of mode choice are based on perceptions rather than objective characteristics, it may be that advertising can play a major role in generating HSR use.

### 3.3 Implications for the Authority's Forecasting Efforts

The above points make a very strong case for continuing model development. The fact that so much of our current understanding of HSR ridership depends upon the calibrated constants means that there are important elements of behavior not explicitly represented in the current model. Some of these can perhaps never be quantified, but it is likely that others can be. A better understanding of trip purposes, their timing, trips as components of longer tours, trip duration and other activities taken during that duration, and traveler reaction to parking conditions or uncertainties about parking availabilities are all examples of factors that might be included in future models. Understanding such factors will also provide benefits to the Authority for future planning needs, such as fare structures, parking provision, and service levels.

The same points argue for continued data collection as the Authority moves toward a more complete understanding of travel behavior as it affects their operations. While the existing data collections efforts are filling important gaps in knowledge, we are quite certain that remaining gaps are also important.

The two-year planning cycle, driven by the Business Plan, has been an obstacle to moving towards a more comprehensive understanding of travel behavior that would diminish the importance of calibrated constants. The Panel urges the Authority to seek ways to promote long-term model development on a separate track from short-term development needed to fulfill the requirements of the Business Plans.

## 4 Socioeconomic Forecasts for Risk Analysis

CS submitted a draft report for review after the meeting dealing with the preparation of the socioeconomic forecasts used in the risk analyses. The report makes several comparisons to forecasts used in the California statewide travel demand model (CSTDM). The analysis and documentation are excellent, and commended as a model for how such inputs should be prepared for all major infrastructure forecasting projects. The amount of judgment involved is acknowledged, but felt appropriate and necessary. A few minor comments are offered:

- CS used the CSTDM forecasts as a starting point, despite believing them to be somewhat extreme. Their defense that this made sense "because they were recently updated to reflect adopted MPO [metropolitan planning organization] forecasts" does not appear compelling.
- Why is the CSTDM projection of job/population ratio used for all three scenarios, given that their forecast of that ratio is said to be "counter to California's trends between WWII and the recent recession?"
- The description of the low-range population forecasts is vague. It is noted that the population "continues the current downward population growth trajectory until reaching 0.5 percent annual growth." The judgment applied to reach that conclusion does not appear to be supported by the time series visualization of the ratio or in the individual forecasts presented in the paper.
- The low and high scenarios, as well as the extreme population distributions, likely have higher probabilities than postulated in Figure S.1 in the draft report. Increasing these probabilities would allay our suspicion that the range appears too narrow. No quantitative reason is given for the probabilities, so they could be changed substantially without any change in the text.

## 5 Conclusions

The calibration and validation work are commendable in their breadth and the highly compressed schedule they were completed within. The briefing book format was effective, and serves as a model for how future briefings to the Panel might be constructed. The Panel has identified a number of areas of concern, but wishes to emphasize their satisfaction with the overall level of effort and results.

Many of the stated concerns stem from limitations of the Versions 1 and 2 modeling systems. As noted in this and prior reports, these can best be addressed by developing a Version 3 modeling system and supporting data collection activities. The Panel will focus further on the requirements for and candidate designs for a Version 3 during their next meeting.